Learning Non-Local Spatial Correlations To Restore Sparse 3D Single-Photon Data

Songmao Chen, *Student Member, IEEE*, Abderrahim Halimi, *Member, IEEE*, Ximing Ren, Aongus McCarthy, *Member, IEEE*, Xiuqin Su, Stephen McLaughlin, *Fellow, IEEE*, and Gerald S. Buller

***Abstract*— This paper presents a new algorithm for the learning of spatial correlation and non-local restoration of single-photon 3D Lidar images acquired in the photon starved regime (fewer or less than one photon per pixel) or with a reduced number of scanned spatial points (pixels). The algorithm alternates between three steps: (i) extract multi-scale information, (ii) build a robust graph of non-local spatial correlations between pixels, and (iii) the restoration of depth and reflectivity images. A non-uniform sampling approach, which assigns larger patches to homogeneous regions and smaller ones to heterogeneous regions, is adopted to reduce the computational cost associated with the graph. The restoration of the 3D images is achieved by minimizing a cost function accounting for the multi-scale information and the non-local spatial correlation between patches. This minimization problem is efficiently solved using the alternating direction method of multipliers (ADMM) that presents fast convergence properties. Various results based on simulated and real Lidar data show the benefits of the proposed algorithm that improves the quality of the estimated depth and reflectivity images, especially in the photon-starved regime or when containing a reduced number of spatial points.**

***Index Terms*— Single photon, 3D Lidar imaging, Poisson statistics, image restoration, ADMM, Laplacian regularization, graph, non-uniform sampling, multi-scale analysis.**

I. Introduction

IMAGING and sensing using the time-correlated single-photon counting technique has emerged as a candidate technology for a number of application areas [1]. The high sensitivity and excellent surface-to-surface resolution of the single-photon timing approach mean that high-resolution depth profiling can be achieved with a very low signal return. This approach has been used in extreme conditions of long stand-off distances [2], [3] and in highly turbid media, as in the case of underwater imaging [4], [5] and atmospheric obscurants [6]. As shown in Fig. 1, a single-photon Lidar system usually consists of a high repetition rate picosecond pulsed laser source which is directed towards the target, with photons returned from the target being detected using a single-photon detector. The time interval between the laser emission and the photon detection event are recorded and digitised into discrete timing intervals. These time intervals are placed in a timing histogram. By recording data over many laser pulses, the timing histogram can represent a full waveform return signal. The integrated sum of photons in the return peak of the waveform represents the target reflectivity, and its time delay infers the round-trip duration, and hence target distance. This allows the reconstruction of 3D image of the target as single-photon measurements are made at different positions on the target.

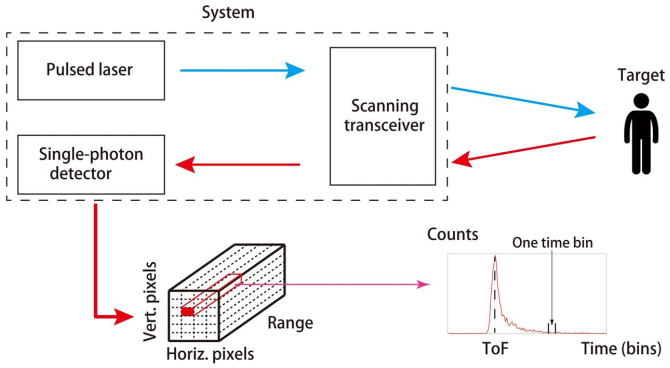

Fig. 1. Schematic description of 3D single photon Lidar imaging.

In order to facilitate the use of single-photon Lidar and imaging systems in a number of emerging applications, it is necessary to address the critical challenges of reducing both the data acquisition time and the image processing computational cost. For example, one possible solution to lowering the acquisition time is to reduce the overall number of acquired pixels by a careful selection of data from parts of the target scene in order to capture most of the useful information [7]. Post-processing the sparse photon image

Manuscript received July 31, 2019; revised November 2, 2019; accepted November 27, 2019. Date of publication December 11, 2019; date of current version January 28, 2020. This work was supported in part by the EPSRC under Grant EP/J015180/1, Grant EP/N003446/1, and Grant EP/M01326X/1 and in part by the U.K. Royal Academy of Engineering through the Research Fellowship Scheme under Grant RF/201718/17128. The work of S. Chen was supported by the China Scholarship Council (CSC) joint Ph.D. Training Program. The associate editor coordinating the review of this manuscript and approving it for publication was Prof. Lucio Marcenaro. *(Corresponding author: Abderrahim Halimi.)*

S. Chen is with the Xi'an Institute of Optics and Precision Mechanics, Chinese Academy of Science (CAS), Xi'an 710119, China, with the University of Chinese Academy of Science (UCAS), Beijing 100049, China, and also with the School of Engineering and Physical Sciences, Heriot-Watt University, Edinburgh EH14 4AS, U.K.

A. Halimi, A. McCarthy, S. McLaughlin, and G. S. Buller are with the School of Engineering and Physical Sciences, Heriot-Watt University, Edinburgh EH14 4AS, U.K. (e-mail: a.halimi@hw.ac.uk).

X. Ren is with the School of Engineering and Physical Sciences, Heriot-Watt University, Edinburgh EH14 4AS, U.K., and also with the Micron School of Materials Science and Engineering, Boise State University, Boise, ID 83725 USA.

X. Su is with the Xi'an Institute of Optics and Precision Mechanics, Chinese Academy of Science (CAS), Xi'an 710119, China, and also with the University of Chinese Academy of Science (UCAS), Beijing 100049, China.

This article has supplementary downloadable material available at http://ieeexplore.ieee.org, provided by the author.

Digital Object Identifier 10.1109/TIP.2019.2957918

data [8] to extract most information is also vitally important, which can include the restoration of 3D images [9], [10], target detection or scene classification [11]. Some of these challenges have been addressed by improving the system optical configurations [12], [13] or by applying signal processing techniques to efficiently exploit sparse single-photon data [14]–[18]. The latter strategy is adopted in this paper.

Several methods have been proposed in the literature to restore Lidar data obtained in the photon-starved regime (fewer or less than one photon per pixel/histogram). As a consequence of the nature of the data, common techniques based on assumptions of Gaussianity lead to a reduced restoration performance which promotes the design of specialized algorithms accounting for data statistics, e.g., binomial statistics [13], Poisson statistics [19]–[25] or data based ones as for machine learning algorithms [8], [26], [27]. The extraction of depth and reflectivity images is an ill-posed inverse problem, that is usually solved by the statistical based algorithms by including prior knowledge or regularization to improve the quality of the estimates. This is often achieved by exploiting redundant information in the data by accounting for correlations or by assuming sparsity in a transformed domain (e.g., DCT or wavelet domain [9], [28]–[30]) to exploit the multi-scale information available in the data. Local approaches assume spatial correlation between adjacent pixels [31], [32] while most state-of-the-art algorithms extend this notion to non-local regions [19], [33], [34]. However, these approaches are no longer effective in the context of Lidar. For example, BM3D [33] fails to restore the empty pixels thus leading to unsatisfactory results as reported in [9]. In addition, larger spatial points such as patches or superpixels [17], [33], [35], [36] can also be used to build robust algorithms and reduce their sensitivity to pixel fluctuations due to noise or other corruptions. The algorithm proposed in this paper combines these ingredients to provide an efficient solution for 3D imaging using Lidar.

The first contribution of this paper is the use of a graph-based strategy to restore 3D Lidar data acquired under the photon starved regime. This choice is motivated by the fact that it allows the exploitation of non-local spatial information available in the data. The possibility of acquiring Lidar data on a non-uniform grid, which requires the use of graph approaches is an additional motivation. Note that any constructed graph should be robust to the photon sparsity inherent in the data. This is achieved by exploiting the multi-scale information of the Lidar data, i.e., the graph nodes are defined by combining the depth and reflectivity information extracted from multi-scale Lidar data. Moreover, to avoid computational difficulties associated with large graphs (e.g., when associating each node to a pixel), the graph-nodes are associated to scene dependent non-uniform patches. More precisely, we propose in this paper to define a non-uniform sampling grid which depends on the observed scene, where larger patches are assigned to homogeneous regions and conversely.

The second contribution exploits the learned non-local spatial correlations to restore depth and reflectivity images in a single-photon Lidar context. Indeed, the use of non-local information is a key ingredient in most state-of-the-art restoration algorithms [19], [33], [34], [23], [24]. Here the restoration is performed by minimizing a regularized cost function coupled with an optimization algorithm. This cost function accounts for the multi-scale information, and promotes non-local spatial correlations between estimates by considering a graph-Laplacian regularization term [37], [38]. The restoration is accelerated by performing a spatial classification on the constructed graph, which allows the independent processing of the patches of each class. The estimates associated with the resulting model are approximated using an alternating direction method of multipliers (ADMM) [39]–[41], as it is fast and shows good convergence properties. The proposed approach is validated on both simulated and field data showing improved depth and reflectivity estimates especially when reducing the photon counts or the number of observed pixels.

To summarize, the main contributions are:

- A graph-based strategy is proposed to learn non-local spatial correlations. The robustness of the graph to low photon regimes is improved using a multi-scale approach, and its computation is accelerated using a non-uniform sampling strategy,
- Exploitation of the learned non-local spatial correlations to restore the depth and reflectivity images. The restoration is accelerated using a clustering strategy allowing parallel processing of the independent classes.

The paper is organized as follows. Section II introduces the statistical model associated with the observed photon counts and an overall description of the proposed solution. Section III presents a preprocessing step based on a multi-scale approach. The spatial correlation analysis and the restoration algorithms are presented in Sections IV and V. The results conducted on synthetic and real data are illustrated in Section VI. Finally, conclusions and future works are discussed in Section VII.

II. Problem Formulation

A. Observation Model

Let $\boldsymbol{y}_{i,j,t}$ be the Lidar observation which denotes the number of observed photon counts within the tth bin of the pixel (i, j), where $(i, j) \in \{1, \cdots, N_r\} \times \{1, \cdots, N_c\}$, and $\boldsymbol{Y}$ be a cube of size $N_r \times N_c \times T$ gathering all observations where T is the total number of bins. Akin to [15], [42], we assume the observed photon counts $\boldsymbol{y}_{i,j,t}$ are drawn from Poisson distribution $\mathcal{P}(.)$ as follow

$$y_{i,j,t} \sim \mathcal{P}\left(s_{i,j,t}\right) \tag{1}$$

where

$$s_{i,j,t} = r_{i,j} f\left(t - t_{i,j}\right) + b_{i,j} \tag{2}$$

and $r_{i,j} \geq 0$ stands for the reflectivity of the target, $t_{i,j} \geq 0$ represents the time-of-flight related to the range of the target, $b_{i,j} \geq 0$ is a constant for all bins denoting the background and dark photon level, and f denotes the system impulse response assumed to be known from the calibration step. Moreover, as $t_{i,j}$ is linearly related to the target depth, it is considered to be depth in this paper. The proposed algorithm aims to

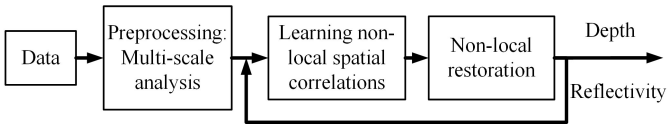

Fig. 2. Overall structure of the proposed algorithm.

restore $r_{i,j}$ and $t_{i,j}$ in the photon starved case by considering the multi-scale information and spatially non-local smoothness constraints obtained on depth and reflectivity (DR) images. Furthermore, let $\boldsymbol{z}$ be a $2N \times 1$ vector where $\boldsymbol{z} = [\boldsymbol{z}_1^T, \boldsymbol{z}_2^T]^T = [\boldsymbol{t}^T, \boldsymbol{r}^T]^T$ gathers all $t_{i,j}, r_{i,j}$ and $N = N_r \times N_c$ is the number of pixels, where both $\boldsymbol{t}$ and $\boldsymbol{r}$ are $N \times 1$ vectors.

B. Structure of Proposed Algorithm

This section introduces the overall structure of the proposed algorithm. As we focus on the photon-starved case, our goal is to use the reflected photons from the target efficiently. Considering a small scanning step compared to the object size, restoration of depth and reflectivity images is usually based on exploiting spatial correlation between pixels. Traditional methods try to find the connection between neighbouring pixels as they are likely to share the same information. However, such methods provide limited improvement due to limited prior information. To improve performance, this paper considers a non-local approach that accounts for spatial correlations in the full image. Based on such idea, we propose a 3 step method which begins by using multi-scale information to obtain initial estimates for DR images. Then it iterates between learning non-local spatial correlations between pixels and using them to restore DR images.

Fig. 2 shows the basic structure of the algorithm, where multi-scale analysis plays the role of a good initialization to guarantee the robustness of the learned correlations. Non-local spatial correlations are then learned using a graph and the algorithm is accelerated using several strategies as detailed in the next section. Finally, the restoration step improves the quality of DR images using the learned correlations.

III. Preprocessing: Multi-Scale Analysis

As the empty and noisy pixels lead to unsatisfactory spatial correlations and thus poor estimation results (see Table I for details), the input data need to be well initialized in order to learn the correct spatial correlations. This section introduces the pre-processing step on DR images which is necessary to improve the quality of the learned non-local spatial correlations. Akin to [17], [43], this papers exploit the information obtained from low-pass filtered histograms which improves the signal-to-background ratio of the data, and provides depth and reflectivity estimates with lower noise at a price of a reduced spatial resolution. In this step, the algorithm first applies a large 2D spatial low-pass filter of size $O \times O$ on the histograms $\boldsymbol{Y}$. The resulting histograms are then used to estimate smooth DR images by considering a classical estimation algorithm based on cross-correlation (see [9] for details). The filter size is then iteratively reduced to obtain DR images with more spatial details, where the empty pixels (i.e. $r_{i,j} = 0$) in finer resolution are inpainted by the corresponding pixels in coarse scale and the noisy pixels are removed by a median filter of size 3×3. The iterations are performed until the filter reaches the finest resolution scale (i.e. $O = 1$). Algo. 1 describes in details these steps, where low-pass($\boldsymbol{Y}$, O) denotes the low-pass filtering of $\boldsymbol{Y}$ with an $O \times O$ filter, inpaint($\boldsymbol{z}^\ell$, $\boldsymbol{z}^{\ell+1}$) replaces the empty pixels in $\boldsymbol{z}^\ell$ by the corresponding pixels in $\boldsymbol{z}^{\ell+1}$, and median($\boldsymbol{z}^\ell$,3) applies a median filter of window size 3×3 on $\boldsymbol{z}^\ell$. The estimation of DR images from histograms using a classical DR estimation algorithm is denoted by Parameter-Estimates(.). Therefore, $\boldsymbol{z}^\ell$ contains the DR images estimated in the ℓ-th iteration from the filtered histograms $\boldsymbol{Y}^\ell$ and the information from previous iteration. The final output $\boldsymbol{z}$ will be the input of the "learning non-local spatial correlations" step as represented in Fig. 2.

Algorithm 1 Multi-Scale Analysis (MSA)

input: $\boldsymbol{Y}$, ℓ^{init}
initialize $\ell \leftarrow \ell^{\text{init}}$, $O \leftarrow 2^{\ell-1}$
$\boldsymbol{z} \leftarrow$ Parameter-Estimates($\boldsymbol{Y}$)
while $O \geq 1$ **do**
  $\boldsymbol{Y}^\ell \leftarrow$ low-pass($\boldsymbol{Y}, O$)
  $\boldsymbol{z}^\ell \leftarrow$ Parameter-Estimates($\boldsymbol{Y}^\ell$)
  if $\ell < \ell^{\text{init}}$ **then**
    $\boldsymbol{z}^\ell \leftarrow$ inpaint($\boldsymbol{z}^\ell$, $\boldsymbol{z}^{\ell+1}$)
    $\boldsymbol{z}^\ell \leftarrow$ median($\boldsymbol{z}^\ell$, 3)
  end if
  $\ell \leftarrow \ell - 1$
  $O \leftarrow 2^{\ell-1}$
end while
Output $\boldsymbol{z} \leftarrow \boldsymbol{z}^\ell$

IV. Spatial Correlations and Acceleration

A. Graph of Correlations

In the non-local approach [38], the spatial correlation between pixels is often evaluated in an $N \times N$ graph $\boldsymbol{W}$ with positive elements ($w_{ij} > 0$) representing the degree of similarity between features/nodes a_i and a_j.[1] Thus the spatial correlation between features can be interpreted as an affinity graph $\boldsymbol{W}$ where the value of each element $w_{i,j}$ is proportional to the similarity of corresponding features. Different heuristic can be chosen to evaluate $\boldsymbol{W}$, for example, Gaussian kernel is applied in this paper as follow [38]

$$w_{i,j} = \exp\left[-\frac{||a_i - a_j||^2}{2\theta^2}\right] \tag{3}$$

where the scaling parameter θ is the kernel's bandwidth, which controls how rapidly the affinity w_{ij} falls with the similarity $||a_i - a_j||^2$ and $||.||$ stands for ℓ_2 norm such that $||\boldsymbol{x}||^2 = \boldsymbol{x}^T\boldsymbol{x}$.

[1]In this context, feature means pixels or patches, which will be described in detail later.

B. Feature

The graph is evaluated using features that should be defined. As we are interested in depth and reflectivity information, both are included as features which will contribute to the robustness of the learned correlations. Furthermore, due to different scales of depth and reflectivity, a normalization is applied on both of them as follows

$$\boldsymbol{t}^{\text{Norm}} = \frac{\boldsymbol{t} - t^{\min}\mathbf{1}}{t^{\max} - t^{\min}}, \quad \boldsymbol{r}^{\text{Norm}} = \frac{\boldsymbol{r} - r^{\min}\mathbf{1}}{r^{\max} - r^{\min}} \tag{4}$$

where $\mathbf{1}$ denotes the $N \times 1$ vector of ones, $\boldsymbol{t}, \boldsymbol{r}$ denotes the input data and $(t^{\max}, r^{\max})$, $(t^{\min}, r^{\min})$ and $(\boldsymbol{t}^{\text{Norm}}, \boldsymbol{r}^{\text{Norm}})$ represent the maximum, minimum and normalized values of $\boldsymbol{t}$ and $\boldsymbol{r}$, respectively.

Let $\alpha \in [0, 1]$ be a parameter which regulates the proportion of depth and reflectivity information, The feature $\boldsymbol{a} \in R^{N\times 1}$ can be written as

$$\boldsymbol{a} = \alpha \boldsymbol{t}^{\text{Norm}} + (1 - \alpha)\boldsymbol{r}^{\text{Norm}}. \tag{5}$$

C. Reducing Complexity

Reducing the computation cost due to the large size of the affinity graph is an essential ingredient to facilitate the use of the method in real world applications. For example, an $N = 128 \times 128$ image contains 16384 pixels, which leads to an affinity graph of size $N \times N = 16384 \times 16384$. This is a large matrix and it is necessary to reduce the computational cost while preserving the correlation between different regions. To solve this problem, a non-uniform sampling strategy is proposed to reduce the size of the graph. Furthermore, the large affinity graph is divided into sub-graphs by adopting a clustering algorithm, which also reduces the complexity.

1) Non-Uniform Sampling: To reduce the size of the graph, down-sampling or grouping of similar pixels in patches is an effective way. For example, a uniform down-sampling with a rate r will lead to an affinity graph of size $[N/(r^2)]\times[N/(r^2)]$. While uniform down-sampling enables computational cost saving, it does not account for the shape of the observed scene. As a result, the image edges and details are degraded while the smoothed area (i.e., homogeneous regions, where average gradient (variation) within the region is lower than the threshold ϵ_s; otherwise it is considered as heterogeneous regions) are assigned too many patches thus resulting in unnecessary computation. From the perspective of saving computational cost, non-uniform sampling is proposed with flexible patch sizes by giving homogeneous regions larger sized patches and smaller sized patches to heterogeneous areas. An example is shown in Fig. 3, where the first row uses uniform sampling and the second row a non-uniformly one. Both images are reconstructed by 1024 patches, which corresponds to 32×32 sized image. It is clear that the non-uniform sampling approach achieves better visual quality for the same number of patches as for a uniformly sampled image. The proposed algorithm is described in Algo. 2 and detailed in the following paragraph. The user needs first to define the coarsest level of resolution L to consider (i.e., patches of size $2^{L-1} \times 2^{L-1}$), the maximum number of patches $N_{\max}$ (i.e. scan points) and the threshold

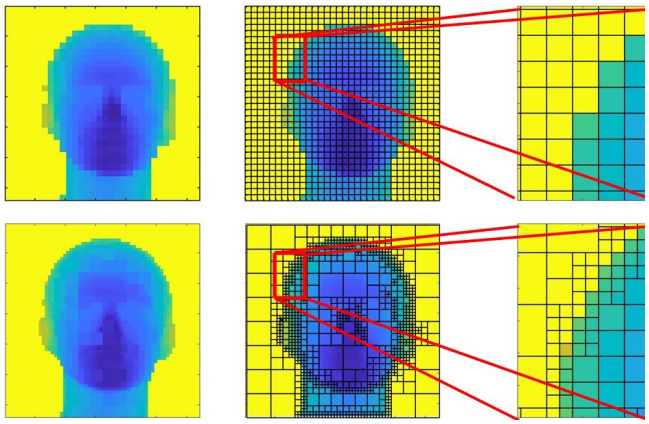

Fig. 3. Sampling results on a life-sized polystyrene head produced using an acquisition time of 30ms per pixel. Top left and bottom left are the results of uniformly and non-uniformly sampled images respectively, top middle and bottom middle are the sampling grid used to obtain top left and bottom left respectively, top right and bottom right represent a zoom on the top left of the head. Both images contain the same number of scan points.

Algorithm 2 Non-Uniform Sampling

1: Input: $\boldsymbol{z}$, L, α, ϵ_T, ϵ_s, $N_{\max}$
2: Initialize: $l \leftarrow L$, $N_p^{'} \leftarrow N/2^{2(L-1)}$
patch sizes $s_n \leftarrow 2^{2(L-1)}, \forall n \in \left\{1, \cdots, N_p^{'}\right\}$
patch locations $\boldsymbol{\nu}$
3: $\boldsymbol{z}^g \leftarrow G(\boldsymbol{z})$
4: $\boldsymbol{z}^{\text{th}} \leftarrow T(\boldsymbol{z}^g, \epsilon_T)$
5: **while** $2^{l-1} > 1$ and $N_p^{'} \leq N_{\max}$ **do**
6: **for** $n = 1 : N_p^{'}$ **do**
7: $c_n \leftarrow \text{Feature}\left(\sum_{i=1}^{s_n} \boldsymbol{z}^{\text{th}}_{\nu_n}(i)/s_n, \alpha\right)$
8: **end for**
9: $\boldsymbol{c} \leftarrow \text{sort}(\boldsymbol{c})$
10: **for** $n = 1 : N_p^{'}$ **do**
11: **if** $c_n > \epsilon_s$ **then**
12: $\boldsymbol{z}^s_n \leftarrow$ split $\boldsymbol{z}_n$
13: **else**
14: $\boldsymbol{z}^s_n \leftarrow$ keep $\boldsymbol{z}_n$
15: **end if**
16: **end for**
17: Update $N_p^{'}, \boldsymbol{s}, \boldsymbol{\nu}, l \leftarrow l - 1$
18: **end while**
19: Output: $\boldsymbol{z}^s, \boldsymbol{\nu}$

to split patches ϵ_T, ϵ_s. Parameter ϵ_T can be fixed based on physical meanings (e.g. a region is considered homogeneous if the variation in depth is lower than 1mm). A binary edge detection is performed in lines 3-4 by first computing the gradients of depth and reflectivity images using $G(.)$, and thresholding the results using ϵ_T and

$$T(z^g, \epsilon_T) = \begin{cases} 1, & \text{if } z^g > \epsilon_T \\ 0, & \text{if } z^g \leq \epsilon_T. \end{cases} \tag{6}$$

The algorithm then iterates from the coarsest patch level to the finest. A weight of variation c_n is associated to each patch as shown in line 7 (where $\boldsymbol{z}^{\text{th}}_{\nu_n}(i)$ is the i-th pixel in the

Algorithm 3 Learning Non-Local Spatial Correlations

1: input $\boldsymbol{z}$
2: initialize L, α, ϵ_T, ϵ_s, $N_{\max}$, K, θ^2
3: $(\boldsymbol{z}^s, \boldsymbol{\nu}) \leftarrow$ Non-uniform sampling($\boldsymbol{z}$, L, α, ϵ_T, ϵ_s, $N_{\max}$)
4: Update N_p: number of patches in $\boldsymbol{z}^s$
5: $\boldsymbol{z}^N \leftarrow$ Normalize($\boldsymbol{z}^s$)
6: $\boldsymbol{a} \leftarrow$ Feature($\boldsymbol{z}^N, \alpha$)
7: **for** $i = 1 : N_p$ **do**
8: **for** $j = 1 : N_p$ **do**
9: $\boldsymbol{w}_{ij} \leftarrow \exp\left[-\frac{||a_i - a_j||^2}{2\theta^2}\right]$
10: **end for**
11: $\boldsymbol{w}_{i,i} \leftarrow 0$
12: **end for**
13: Label $\leftarrow$ clustering($\boldsymbol{W}, K$), e.g., using [48]
14: Output $\boldsymbol{z}^s, \boldsymbol{\nu}$, Label, $\boldsymbol{W}$

n-th patch). Note that feature(.) in line 7 combines the depth and reflectivity information to build a feature as described in (5). The weights are then sorted in decreasing order in line 9, so that priority is given to those patches with high variations. Finally, patches with weights higher than ϵ_s are split in lines 10-16 (smooth patches are not divided). The algorithm is stopped if it reaches the finest level, or the maximum number of patches $N_{\max}$. Finally, the outputs $\boldsymbol{z}^s$ and $\boldsymbol{\nu}$ contain the patch values and locations, respectively.

2) Clustering: The affinity graph $\boldsymbol{W}$ is constructed using the non-uniformly sampled patches as features. Clustering is then performed to divide the large affinity graph $\boldsymbol{W}$ into smaller sub-graphs that can be processed independently, thus leads to better computational performance (see also Table I). Several graph-based clustering algorithms can be considered [44]–[47] and we adopt in this paper the clustering algorithm in [48] that shows a good balance between computational cost and clustering quality. This algorithm only requires the affinity graph $\boldsymbol{W}$ and the number of classes K to divide the patches into K classes, where the sub-graphs $\boldsymbol{W}^k, k \in \{1, \cdots, K\}$ associated with each class can be obtained by selecting the appropriate lines and columns of $\boldsymbol{W}$. The algorithm performs an SVD decomposition on $\boldsymbol{W}$, then apply K-means algorithm on the eigenvectors matrix of $\boldsymbol{W}$. Although [47] generalizes [48] by using several iterations, [48] is considered in this paper as it provided satisfactory results at a low computational cost. The learning non-local spatial correlations step, including non-uniform sampling, is described in Algo. 3. Note that line-3 is described in Algo. 2. Normalize(.) and Feature(.) correspond to (4) and (5), respectively. To summarize, this step outputs the non-uniformly sampled DR images $\boldsymbol{z}^s$, the patch locations $\boldsymbol{\nu}$, the class labels of the patches and the affinity graph $\boldsymbol{W}$, which will be the input of the restoration step described in the next section.

V. Restoration

The goal of the paper is to estimate the depth $\boldsymbol{t}$ and reflectivity images $\boldsymbol{r}$ in the photon starved regime, where the main challenge is the presence of empty and noisy pixels due to the limited photon counts. From an optimization perspective, such goal can be reached by accounting for prior knowledge introduced by regularization terms. The next section describes the considered cost function to minimize.

A. Cost Function

The cost function, denoted by $\mathcal{C}(\boldsymbol{t}, \boldsymbol{r})$, is given by

$$\mathcal{C}(\boldsymbol{t}, \boldsymbol{r}) = \rho(\boldsymbol{t}, \boldsymbol{r} | \tilde{\boldsymbol{t}}, \tilde{\boldsymbol{r}}) + \phi(\boldsymbol{t}, \boldsymbol{r}) + i_{\mathbb{R}_+}(\boldsymbol{t}, \boldsymbol{r}) \tag{7}$$

where $\rho(.)$ is the data fidelity term using non-uniform patches, $\phi(.)$ is the regularization term representing a negative log-prior distribution on the parameters $(\boldsymbol{t}, \boldsymbol{r})$ and $i_{\mathbb{R}_+}(.)$ is the indicator function imposing non-negativity. These terms are described in the next sub-sections.

1) Fidelity Term: To estimate spatial correlations effectively, priority was given to remove empty and noisy pixels at the preprocessing step, leading to a smooth depth image $\boldsymbol{t}^s$ with reduced spatial details. To retrieve these spatial details, a new depth image is reconstructed using the following equation

$$\hat{\boldsymbol{t}} = \boldsymbol{M}\boldsymbol{t}^y + (\mathbb{I}_N - \boldsymbol{M})\boldsymbol{t}^s \tag{8}$$

where $\boldsymbol{t}^y$ represents the estimated depth from the fine scale histograms $\boldsymbol{Y}$, and $\boldsymbol{M}$ is an $N \times N$ diagonal binary mask, with ones corresponding to non-empty pixels in $\boldsymbol{t}^y$. This operation replaces empty pixels in $\boldsymbol{t}^y$ by the corresponding ones in $\boldsymbol{t}^s$ to keep spatial details. Note that due to inherent Poisson noise in the reflectivity image, it is not changed at this stage leading to $\hat{\boldsymbol{r}} = \boldsymbol{r}^s$.

Furthermore, it is important to mention that the estimation is performed on the non-uniform patches, thus the pixel-wise images $(\hat{\boldsymbol{t}}, \hat{\boldsymbol{r}})$ are used to build the patch-wise images $(\tilde{\boldsymbol{t}}, \tilde{\boldsymbol{r}})$, where the value of each patch is given by the median value of its pixels. Finally, by assuming a Gaussian distribution of the estimates, the data fidelity term is given by

$$\rho(\boldsymbol{t}, \boldsymbol{r} | \tilde{\boldsymbol{t}}, \tilde{\boldsymbol{r}}) = ||\boldsymbol{t} - \tilde{\boldsymbol{t}}||^2 + ||\boldsymbol{r} - \tilde{\boldsymbol{r}}||^2 \tag{9}$$

Note that the fidelity term is separable w.r.t $\boldsymbol{r}$ and $\boldsymbol{t}$ which can be processed independently. Such advantages will be illustrated in subsequent sections.

2) Prior: Prior knowledge plays a key role in restoring the estimates of empty and noisy pixels [49]. Note that the correlations between pixels have been learned in the previous step, which constitutes valuable information that should be exploited for restoration. Akin to [50], we assume a centred multivariate Gaussian prior distribution for the pixels of depth $\boldsymbol{t}_k$ and reflectivity $\boldsymbol{r}_k$ images belonging to the spatial class k, where the correlations are introduced through the covariance matrices. Straightforward computations lead to the following negative log-prior or Laplacian regularization

$$\phi(\boldsymbol{t}_k, \boldsymbol{r}_k) = \sigma_1 \boldsymbol{t}_k^T \boldsymbol{P}^k \boldsymbol{t}_k + \sigma_2 \boldsymbol{r}_k^T \boldsymbol{P}^k \boldsymbol{r}_k \tag{10}$$

where σ_1 and σ_2 control the degree of regularization, $\boldsymbol{P}^k = \boldsymbol{Q}^k - \boldsymbol{W}^k$ is the graph Laplacian matrix of the k-th class in which $\boldsymbol{Q}^k_{ii} = \sum_{j=1}^N w^k_{ij}$ is a diagonal matrix and $\boldsymbol{W}^k$ is the affinity graph of the pixels belonging to the k-th class. By combining the fidelity term and negative log-prior distributions for the k-th spatial class, it is clear that $\mathcal{C}(\boldsymbol{t}_k, \boldsymbol{r}_k)$

is separable w.r.t $\boldsymbol{t}$ and $\boldsymbol{r}$, which is divided into the following 2 similar sub-functions that are optimized independently for better efficiency.

$$\begin{aligned}\mathcal{C}(\boldsymbol{t}_k) &= ||\boldsymbol{t}_k - \tilde{\boldsymbol{t}_k}||^2 + \sigma_1 \boldsymbol{t}_k^T \boldsymbol{P}^k \boldsymbol{t}_k + i_{\mathbb{R}_+}(\boldsymbol{t}_k) \\ \mathcal{C}(\boldsymbol{r}_k) &= ||\boldsymbol{r}_k - \tilde{\boldsymbol{r}_k}||^2 + \sigma_2 \boldsymbol{r}_k^T \boldsymbol{P}^k \boldsymbol{r}_k + i_{\mathbb{R}_+}(\boldsymbol{r}_k) \end{aligned} \quad (11)$$

Finally, it is clear from (11), that each class k can be restored independently from the others. Therefore, we drop the index "k" in the following for brevity.

B. Estimation Algorithm

This section describes the alternating direction method of multipliers (ADMM) [51] used to minimize the cost function (11). The main idea of ADMM algorithm is to divide the formulation into a sets of simple sub problems that are easy to solve. This section first introduces a generalized formulation of a variant of the ADMM algorithm [41], then demonstrates its use in our case.

1) General Formulation: Consider the generalized optimization problem which contains J sub problems

$$\underset{\boldsymbol{z}}{\operatorname{argmin}} \mathcal{C}(\boldsymbol{z}) = \underset{\boldsymbol{z}}{\operatorname{argmin}} \sum_{j=1}^{J} g_j\left(\boldsymbol{H}^{(j)}\boldsymbol{z}\right) \quad (12)$$

where $\boldsymbol{z} \in \mathbb{R}^d$, $g_j : \mathbb{R}^{p_j} \to \mathbb{R}$ are closed, proper, convex functions, and $\boldsymbol{H}^{(j)} \in \mathbb{R}^{p_j \times d}$ are arbitrary matrices. By denoting $\boldsymbol{u}^{(j)} = \boldsymbol{H}^{(j)}\boldsymbol{z} \in \mathbb{R}^{p_j}$ and $\boldsymbol{d}^{(j)} \in \mathbb{R}^{p_j}$ which is introduced as an auxiliary vector, problem (12) can be solved using a variant of the ADMM algorithm proposed in [40] and described in Algo. 4. The algorithm first reforms (12) in augmented Lagrangian form as follows

$$\begin{aligned}&\underset{\boldsymbol{z},\boldsymbol{u}}{\operatorname{argmin}} \mathcal{C}(\boldsymbol{z},\boldsymbol{u}) \\ &= \underset{\boldsymbol{z},\boldsymbol{u}}{\operatorname{argmin}} \sum_{j=1}^{J}\left[g_j\left(\boldsymbol{u}^{(j)}\right) + \frac{\mu}{2}||\boldsymbol{u}^{(j)} - \boldsymbol{H}^{(j)}\boldsymbol{z} - \boldsymbol{d}^{(j)}||^2\right]. \end{aligned} \quad (13)$$

Then it alternates between the minimization of (13) w.r.t. $\boldsymbol{z}$ and $\boldsymbol{u}$. Note that μ stands for the Lagrange multiplier that has been updated using the adaptive strategy described in [52] to accelerate convergence; g_j are closed, proper, convex functions, and the proximity term in line 11 are solved exactly. This algorithm convergence is guaranteed when the matrix $\boldsymbol{X} = \left[\sum_{j=1}^{J}\left(\boldsymbol{H}^{(j)}\right)^\top \boldsymbol{H}^{(j)}\right]$ has full rank and ϵ_c is the threshold determining whether the algorithm has converged

2) Restoration : Problem (11) can be solved using the ADMM algorithm. Denoting $\boldsymbol{t}$ (resp. $\boldsymbol{r}$) a vector of size $N_p^k \times 1$, the sub-function $\mathcal{C}(\boldsymbol{t})$ (resp. $\mathcal{C}(\boldsymbol{r})$) can be divided into 2 sub-problems as follows

$$\begin{aligned} g_1\left(\boldsymbol{u}^{(1)}\right) &= ||\boldsymbol{u}^{(1)} - \tilde{\boldsymbol{t}}||^2 + \sigma_1 \boldsymbol{u}^{(1)T}\boldsymbol{P}\boldsymbol{u}^{(1)}, \quad \boldsymbol{H}^{(1)} = \mathbb{I}_{N_p^k} \\ g_2\left(\boldsymbol{u}^{(2)}\right) &= i_{\mathbb{R}_+^{N_p}}\left(\boldsymbol{u}^{(2)}\right), \quad \boldsymbol{H}^{(2)} = \mathbb{I}_{N_p^k} \end{aligned} \quad (14)$$

where $\boldsymbol{X}$ simplifies to a variant of identity matrix (which is fast to inverse) due to the unique structure of $\boldsymbol{H}^{(j)}$.

Algorithm 4 ADMM Variant for (12)

1: input: ϵ_c, g_j, $\boldsymbol{H}^{(j)}, \forall j$
2: Initialize $\boldsymbol{z}, \boldsymbol{u}_0^{(j)}, \boldsymbol{d}_0^{(j)}, \forall j$, μ, $i_{\max}$
3: Set $i \leftarrow 0$, conv$\leftarrow 0$
4: **while** conv$= 0$ **do**
5: **for** j=1:J **do**
6: $\lambda_i^{(j)} \leftarrow \boldsymbol{u}_i^{(j)} + \boldsymbol{d}_i^{(j)}$,
7: **end for**
8: $\boldsymbol{z}_{i+1} \leftarrow \boldsymbol{X}^{-1}\sum_{j=1}^{J}\left(\boldsymbol{H}^{(j)}\right)^\top \lambda_i^{(j)}$,
9: **for** j=1:J **do**
10: $\boldsymbol{v}_i^{(j)} \leftarrow \boldsymbol{H}^{(j)}\boldsymbol{z}_{i+1} - \boldsymbol{d}_i^{(j)}$,
11: $\boldsymbol{u}_{i+1}^{(j)} \leftarrow \underset{\boldsymbol{s}}{\operatorname{argmin}} \frac{\mu}{2}||\boldsymbol{s} - \boldsymbol{v}_i^{(j)}||^2 + g_j(\boldsymbol{s})$,
12: $\boldsymbol{d}_{i+1}^{(j)} \leftarrow \boldsymbol{d}_i^{(j)} - \left(\boldsymbol{H}^{(j)}\boldsymbol{z}_{i+1} - \boldsymbol{u}_{i+1}^{(j)}\right)$,
13: **end for**
14: **if** $||\boldsymbol{z}_{i+1} - \boldsymbol{z}_i|| < \epsilon_c$ **or** $(i > i_{\max})$ **then**
15: conv $\leftarrow 1$
16: **end if**
17: $i = i + 1$
18: **end while**
19: output: $\boldsymbol{z}_{i+1}$

Algorithm 5 Non-Local Restoration

1: Input: $\boldsymbol{Y}$, $\boldsymbol{z}^s$, $\boldsymbol{z}^y$, Label, $\boldsymbol{W}$
2: Initialize $\tilde{\boldsymbol{t}}$, $\tilde{\boldsymbol{r}}$, σ_1, σ_2, $\boldsymbol{M}$
3: **for** k=1:K **do**
4: $\boldsymbol{W}^k$ = class($\boldsymbol{W}$, Label)
5: Compute $\boldsymbol{Q}^k$
6: $\boldsymbol{P}^k = \boldsymbol{Q}^k - \boldsymbol{W}^k$
7: $\phi(\boldsymbol{t}_k, \boldsymbol{r}_k) = \sigma_1 \boldsymbol{t}_k^T \boldsymbol{P}^k \boldsymbol{t}_k + \sigma_2 \boldsymbol{r}_k^T \boldsymbol{P}^k \boldsymbol{r}_k$
8: $(\boldsymbol{t}_k, \boldsymbol{r}_k) \leftarrow \underset{\boldsymbol{t}_k, \boldsymbol{r}_k}{\operatorname{argmin}} \mathcal{C}(\boldsymbol{t}_k, \boldsymbol{r}_k)$ using Algo. 4
9: **end for**
10: Output: $\boldsymbol{z} = [\boldsymbol{t}^T; \boldsymbol{r}^T]^T$

The optimization problem in line 11 is straightforward for $j = 1$ with the following analytical solution

$$\boldsymbol{u}_{i+1}^{(1)} = \left[(\mu + 2)\mathbb{I}_{N_p} + \sigma_1\left(\boldsymbol{P} + \boldsymbol{P}^T\right)\right]^{-1}(\mu \boldsymbol{v}_i^{(1)} + 2\tilde{\boldsymbol{t}}). \quad (15)$$

The restoration step is described in Algo. 5, where $(.)^k$ means the parameters in k-th class, class($\boldsymbol{W}$, Label) stands for the choice of the patches of the k-th class. Finally, as highlighted in section II-B, the proposed algorithm iterates between learning non-local spatial correlations and non-local restoration steps to improve the quality of DR images.

VI. Simulation Results

A. Results on Synthetic Data

This section evaluates the performance of the proposed algorithm on synthetic data sets. It is divided into four subsections which introduce the evaluation criteria and comparison algorithms, the considered data sets, a study of the effect of hyperparameters, and analysis of performance on simulated images, as follows.

1) Evaluation Criteria and Comparison Algorithms: The quality of restoration is evaluated using the Signal to Reconstruction-Error ratio, $\text{SRE} = 10\log_{10}\left(\frac{||\boldsymbol{x}||^2}{||\boldsymbol{x}-\widehat{\boldsymbol{x}}||^2}\right)$ and Root Mean-Square Error, $\text{RSME} = \sqrt{\frac{||\boldsymbol{x}-\widehat{\boldsymbol{x}}||^2}{N}}$, where $\boldsymbol{x}$ is the reference depth or reflectivity image, $\widehat{\boldsymbol{x}}$ is the restored image. For synthetic data, the reference images are the known parameters used to generate the data. For real data, DR images obtained with the longest acquisition times were considered as references.

Different methods are considered for comparison. A classical method is considered by performing a pixel-wise matched filter of the histograms by the instrument response of the system to estimate the depth (this is equivalent to a maximum likelihood estimator by considering Gaussian statistics and no background). Both Restoration of Depth and Intensity-Total Variation (RDI-TV) [9] and Shin's algorithm [53] restore DR images by considering data statistics and local spatial correlations enforced using a TV regularization term. Unmixing Algorithm (UA) [17] improves the quality of the estimate by considering a data-driven adaptive super-pixel approach, then further improves the performance of restoration by solving a penalized maximum likelihood estimates that account for data statistics and spatial correlations. All 4 algorithms were implemented in MATLAB 2018a on a PC of Intel i7-9700k OCTA Core CPU and 32G RAM.

2) Simulated Data Sets: We test the algorithm on two synthetic data obtained from simulated and realistic images. The first data consists of a $128 \times 128 \times 586$ ($N_r \times N_c \times T$) sized synthetic data representing a panda face scene. The Lidar observation of each pixel $y_{i,j,t}$ is generated by the observation model in (2) and the reference depth and reflectivity images are shown in Fig. 4 (left), where the ears, eyes and noise of the panda face share the same reflectivity value, while the remaining regions share another value. For the depth, the face is spherical with a bump on the nose. To further test the proposed algorithm in complex scenarios, a second real scene from Middlebury dataset [54] is used to simulate photon count histograms. More precisely, $192 \times 288 \times 586$ ($N_r \times N_c \times T$) sized image of a recycle with curtain and a board is used in the experiment as can be seen in Fig. 5. For both scenarios, the number of received photons is varied to simulate different acquisition times, which is evaluated by the average number of photon-per-pixel (PPP). To satisfy the assumption that the noise $\boldsymbol{b}$ is neglectable, we consider a small value of $\boldsymbol{b} = 0.0001$ leading to the SBR values 58.76, 232.44, 1153.37 and recycle scene are 34.29, 167.77, 333.97 respectively.

3) Effect of Hyperparameters: This section highlights the benefit of each step of the proposed algorithm and describes the selection criteria of the hyperparameters. The maximum size of the low-pass filter O depends on the PPP level in the observed data [17], where a larger sized filter should be applied for low PPP. In this work, we fix this parameter using:

$$\underset{O}{\text{argmin}}[(\text{PPP} \times O^2 - y_s) > 0] \tag{16}$$

where $O = 2^{(\lambda-1)}$ and λ is a positive integer, y_s is a fixed desirable number of photons in each superpixel, e.g. 50.

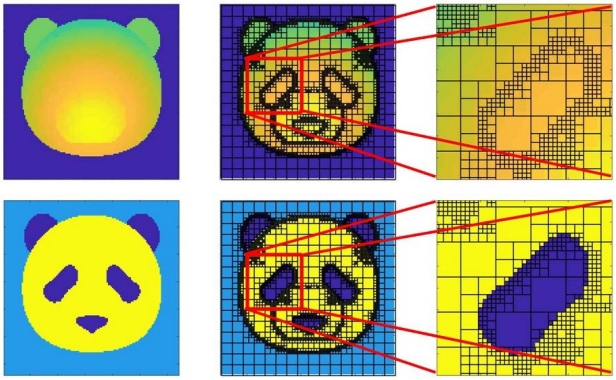

Fig. 4. Patches obtained by non-uniform sampling. The top row represents depth images and the bottom row is the reflectivity images. Images from left to right are sampled reference images, the non-uniform grid and the zoomed-in image of the left eye.

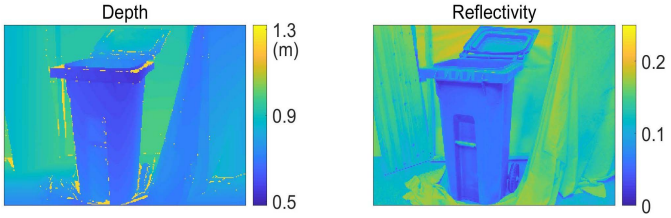

Fig. 5. Reference images of the Recycle dataset extracted from the Middlebury database [54].

In addition, note that most of the parameters of the proposed algorithm can be related to physical meaning (e.g., the thresholds for patch division ϵ_T) or driven by data. For example, the choice of α and K is based on our prior knowledge of the scene while σ_1 and σ_2 are proportional to the variance of each class. We illustrate the effect of some hyperparameters using an example. Table I shows the performance of the proposed algorithm when varying the maximum size $O^{\max}$ of the low-pass filter (the considered scale) and the number of spatial classes K on the panda scene with PPP $= 0.8$. From this table, it is clear that considering multi-scale information is essential to improve performance as reflected by the SRE results of different filter sizes. A small filter size leads to poor spatial correlations while a larger filter can degrade the fine details. Moreover, the table shows that the number of classes affects the computational cost, where a higher number leads to lower cost. However, the SRE are only robust if the number of classes is in a reasonable range. Note that $K = 12$ is fixed when varying $O^{\max}$, similarly, $O^{\max} = 8$ is also fixed when testing different K.

4) Performance on Simulated Data: This section studies the performance of the proposed algorithm on simulated data. The ADMM algorithm shows fast convergence and we fix the maximum number of iterations to 50 for all experiments. Considering the recycle scene for example, Fig. 6 justifies the considered maximum iterations. Akin to Fig. 2, the loop between the "learning non-local spatial correlations" step and "non-local restoration" step is stopped when the average absolute difference between the estimated images and previous estimated images satisfies depth $<$ 5 time bin (i.e., 1.5mm

TABLE I

PERFORMANCE OF THE PROPOSED ALGORITHM WITH RESPECT TO DIFFERENT HYPERPARAMETERS. THE RESULTS ARE OBTAINED ON THE PANDA FACE WITH PPP = 0.80

Maximum Filter Size ($O^{\max}$)	32	16	8	4	2	1
SRE of Depth	20.3	20.2	20.4	20.1	17.4	3.1
SRE of Reflectivity	11.4	11.7	11.8	10.2	7.0	1.2
Number of Classes (K)	24	12	8	4	2	1
SRE of Depth	20.4	20.4	20.5	20.3	20.2	20.1
SRE of Reflectivity	11.7	11.8	11.7	11.7	11.8	11.6
Processing time	3.3	3.6	3.7	7.7	15.1	37.3

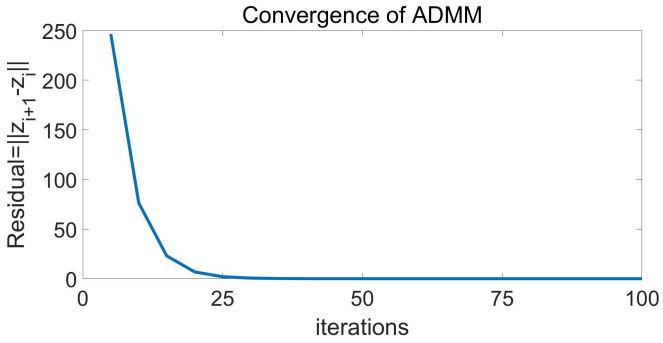

Fig. 6. An example of the ADMM convergence, where x-axis represents iterations and y-axis corresponds to the norm between successive estimates.

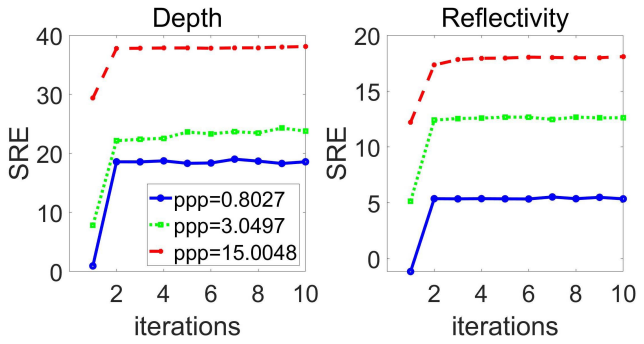

Fig. 7. SRE of depth (left) and reflectivity (right) on panda face with iterations.

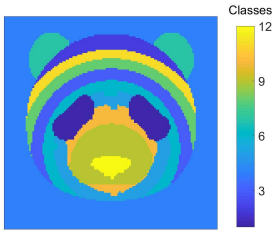

Fig. 8. Classification result on the reference image of panda face (12 classes).

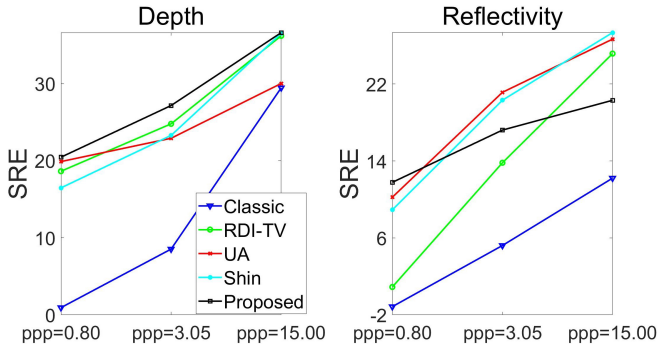

Fig. 9. SRE comparison of depth (left) and reflectivity (right) of different methods on panda scenario. The proposed algorithm uses 8192 non-local patches while other methods uses full resolution (i.e., 128 × 128 pixels).

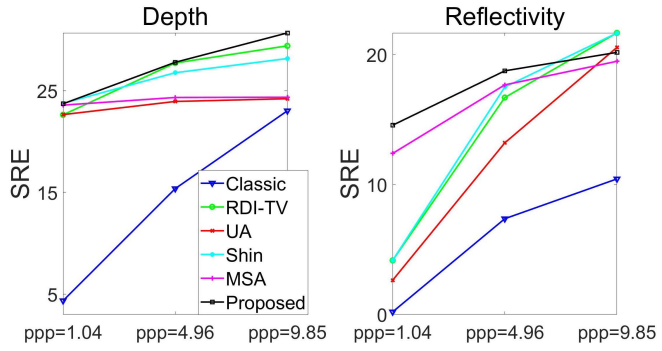

Fig. 10. SRE comparison of depth (left) and reflectivity (right) of different methods on the recycle scenario. The proposed algorithm uses 13824 non-uniform patches while other algorithms accessed the full resolution.

by considering bin width of 2ps) and reflectivity $< 20\%$. Note that the DR images from the multi-scale step are considered as the first iteration of estimated image. In the following, the quantitative results have been obtained while considering 10 Monte-Carlo simulations. The hyperparameters of the proposed algorithm are set to be $\theta = 100$, $L = 5$, $K = 24$, $\alpha = 0.95$ and $N_{\max} = 8192$ for all the panda face experiments while O is 8, 4 and 2 for PPP= 0.80, PPP= 3.05, PPP= 15.00 cases. Another example of non-uniform sampling is represented in Fig. 4. More importantly, by analysing the right column it appears that the non-uniform sampling is achieved using information from both DR images thanks to the feature step illustrated in section IV-B. In the right column, the edge of the eye is detected by the reflectivity which varies in this region, while the depth is almost constant. On the other hand, the small patches located at the bottom right of the eye are mainly due to depth variation in this region. Thanks to this non-uniform sampling, the algorithm identifies each class robustly as some classes share either similar depth or reflectivity (e.g. face and eyes). A classification example is shown in Fig. 8, where the panda face is divided into 12 classes. The patches of each class can be non-local allowing the exploitation of non-local similarities for DR image restoration.

The comparison of different methods using SRE is shown in Fig. 9 where the algorithm proposed in this paper is applied using a limited number of patches while other algorithms have access to the full resolution image. All of the algorithms perform better than the classical algorithm for both depth and reflectivity estimation. However, the proposed algorithm shows better performance than UA and Shin's algorithm for the depth results. Regarding to reflectivity, the proposed algorithm shows best performance under the photon starved case while both UA and Shin's algorithm perform better for higher photon count cases. The processing time of different algorithms are investigated in columns 2, 3 and 4 in Table II

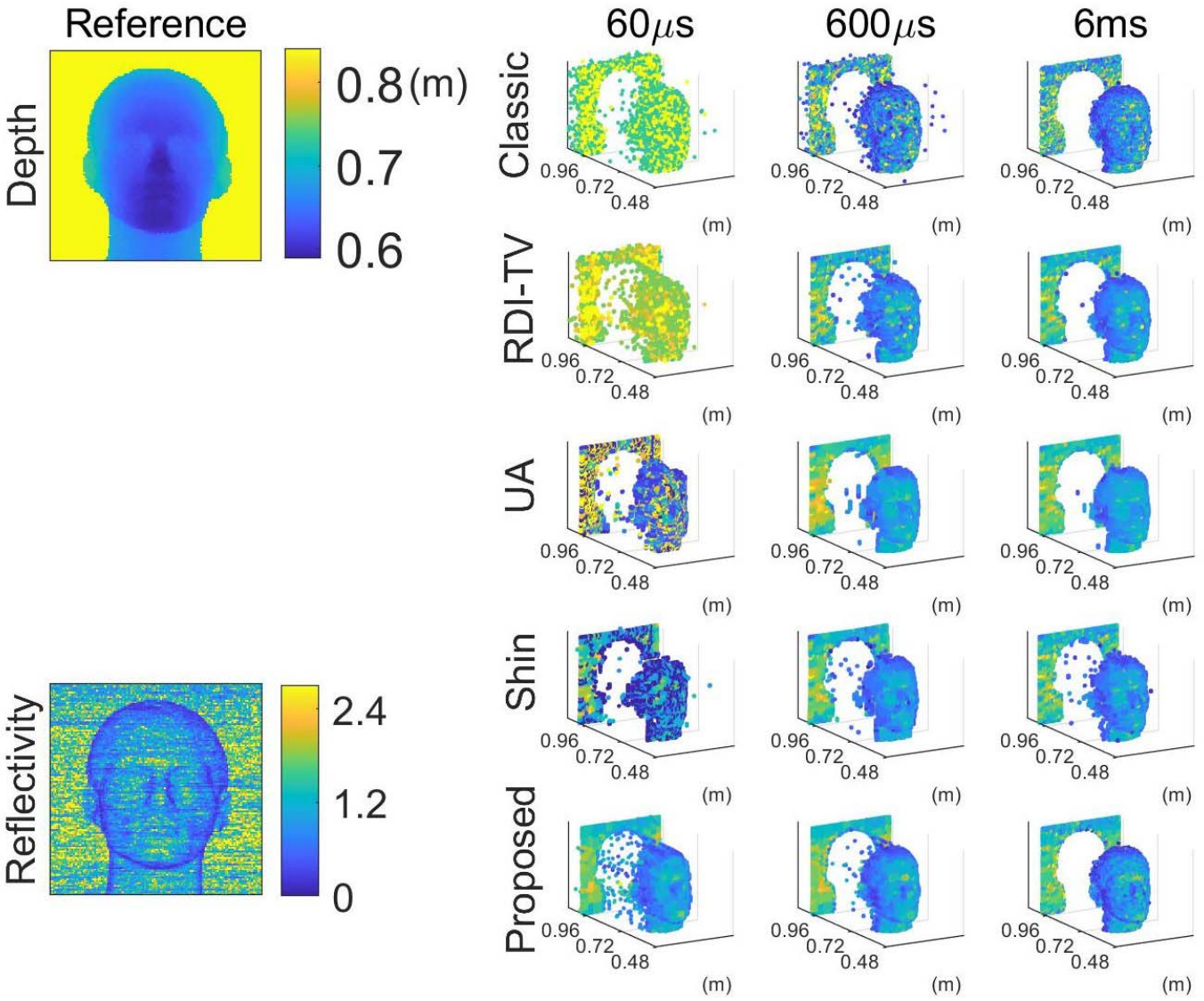

Fig. 11. Depth and reflectivity reconstruction of mannequin face. The first column is the reference images (30ms image). The second to fourth column corresponds to different acquisition times, i.e., 60μs, 600μs, 6ms and the rows correspond to different algorithms. Reflectivity is color coded.

TABLE II
PROCESSING TIME (IN SECONDS) OF THE ALGORITHMS COMPARED ON PANDA SCENARIO (LEFT 3 COLUMNS) AND RECYCLE SCENARIO (RIGHT 3 COLUMN)

	Panda			Recycle		
PPP	0.80	3.05	15.00	1.04	4.96	9.85
Proposed	**3.08**	**2.86**	**2.52**	**8.26**	**7.31**	**7.22**
RDI-TV[9]	8.31	5.15	3.55	15.39	14.12	11.02
Shin[53]	6.09	5.46	4.23	16.12	12.22	8.67
UA[17]	8.57	6.07	4.20	12.87	9.82	8.82

where the proposed algorithm works under 8192 non-uniform patches and other methods process the full resolution image (128×128). Both UA and Shin's algorithm adopted parallel computing strategy which justifies their fast computations.

To confirm these results, the proposed algorithm was also applied to the recycle scene while considering three PPP levels and the following hyperparameters: $\theta = 100$, $L = 6$, $K = 20$, $\alpha = 0.95$. $N_{\max} = 13824$, and O is 8, 4 and 4 for PPP equals to 1.04, 4.96, and 9.85. Numerical results are shown in Fig. 10 where all the algorithms improve the SRE significantly both in depth and reflectivity with respect to the classical algorithm. This figure includes the results of the initialization step described in Algo. 1 and denoted by MSA (Multi-Scale Analysis). The benefit of the non-local restoration step is highlighted by the clear improvement obtained between the MSA and the proposed algorithm. RDI-TV, UA and Shin's algorithms show similar performance while the proposed algorithm shows best performance, especially on the reflectivity result in the sparse photon case. Finally, the processing times of the algorithms on the recycle scene are listed in columns 5, 6 and 7 of Table. II.

B. Results on Real Data

This section considers six real images (128×128 pixels and 586 time bins) of a life-sized white polystyrene head, acquired at a distance of 40m. The images were

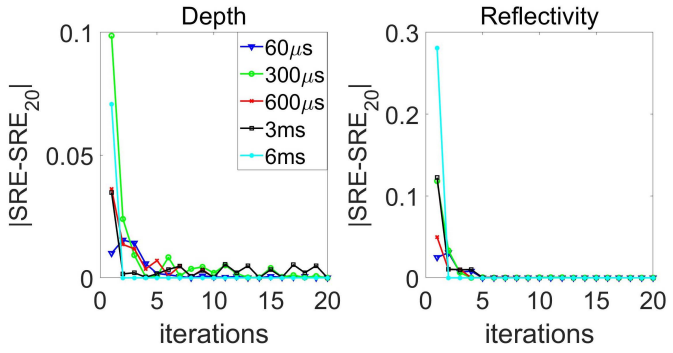

Fig. 12. Algorithm's convergence with iterations, where the X-axis is the iterations and the Y-axis is the average of absolute difference of SRE between k-th estimated image and the 20-th estimated DR images.

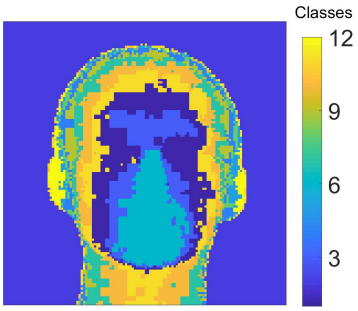

Fig. 13. Classification result obtained on 600μs data (12 classes).

acquired in November 2014 on the Edinburgh Campus of Heriot-Watt University, using a time-of-flight scanning sensor, based on time-correlated single-photon counting (TCSPC) technique [12]. The transceiver system and data acquisition hardware used for this work are broadly similar to that described in [12], [14] (see also [15] for more details regarding the system parameters). A time bin width of 16ps was used for the TCSPC data acquisition. The acquired data contains time tagged detection, thus, it is possible to simulate a reduction in acquisition time. In this paper, we consider images having the following acquisition time per pixel, 60μs, 300μs, 600μs, 3ms, 6ms and 30ms, which correspond to PPP levels of 0.80, 4.09, 8.21, 41.06, 82.02, 410.00. The proposed algorithm is run using $\theta = 100$, $L = 5$, $K = 12$, $\alpha = 1$ and $N_{\max} = 4096$ for all the mannequin face experiments, O is 8, 4, 2, 2 and 1 for the five scenarios. Moreover, all other methods operate on uniformly down-sampled image of size $64 \times 64 \times 586$, i.e., 4096 pixels in all experiments. Similar to the process of recycle, the restored downsampled DR images obtained by other algorithms are upsampled to full resolution by sharing the same information for each 2 pixels to compute SRE and exhibition.

Fig. 11 shows good restoration of 3D images for 60μs, 600μs, 6ms, especially for the sparse case where empty and noisy pixels are frequent. Fig. 12 illustrates the good convergence of the proposed algorithm where the results obtained in each iteration are compared with estimated images after 20 iterations, i.e., $\boldsymbol{t}^{20}$ and $\boldsymbol{r}^{20}$ respectively. This figure shows good convergence behaviour where all depth and reflectivity curves converge after few iterations. The classification result performed on 600μs data is shown in Fig. 13. The image

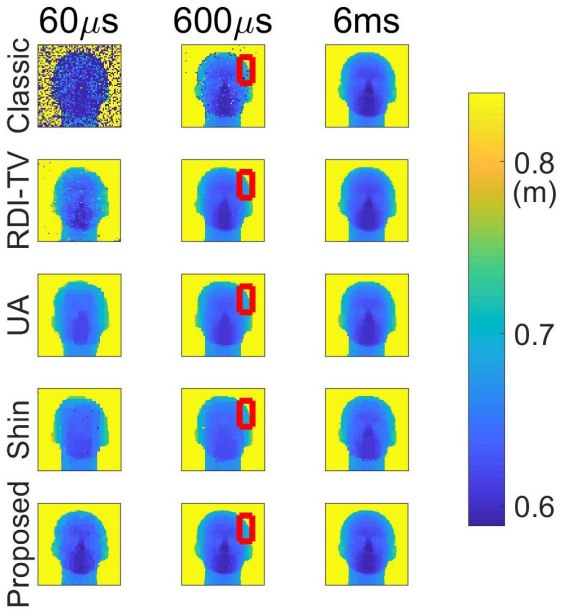

Fig. 14. Depth maps obtained by different methods on mannequin face. Columns from left to right corresponds to data of different acquisition times, i.e., 60μs, 600μs, and 6ms. Rows from top to bottom corresponds to different methods, i.e., cross-correlation, RDI-TV, UA, Shin's algorithm and proposed algorithm.

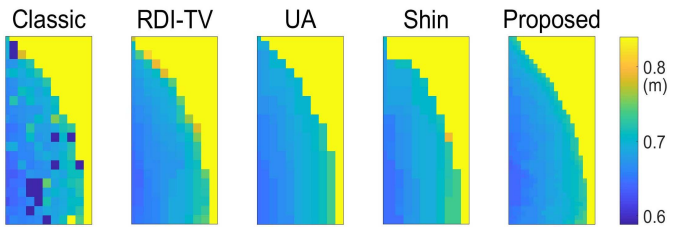

Fig. 15. Zoomed-in image of head on the 600μs images obtained by different algorithms (red blocks in Fig. 14).

is classified into 12 classes where 4096 patches were used by considering non-uniform sampling. The obtained regions share similar depth information, and highlights the features of the mannequin face. Fig. 14 shows the depth results obtained by different methods. All algorithms performed better than the classical cross-correlation algorithm with a clear advantage shown by the proposed algorithm. This highlights the effectiveness of the non-uniform sampling as confirmed in Fig. 15 showing the zoomed in image of the red blocks in Fig. 14.

The restoration of reflectivity is shown in Fig 16, which is considering 4096 patches. Thanks to the learned non-local correlations, the proposed algorithm shows the best results for all 60μs, 600μs and 6ms data by preserving the face's shape and reducing the noise. RDI-TV, UA and Shin's algorithm shows similar performance in 600μs and 6ms data, which smooth the image but key information is also degraded. The proposed algorithm is the only one showing acceptable results at 60μs, which highlights its superior performance under the photon starved regime.

In order to compare the results obtained by the different methods numerically, SRE is used to evaluate their performance. The reference image is the 30ms image for both depth

TABLE III
RSME OF DEPTH RESULTS OBTAINED BY DIFFERENT METHODS (IN METRES)

	Panda			Recycle			Mannequin face				
PPP	0.80	3.05	15.00	1.04	4.96	9.85	0.80	4.09	8.21	41.06	82.02
Classic[9]	0.1783	0.0744	0.0067	0.4907	0.1380	0.0572	0.5641	0.2147	0.1028	0.0157	0.0143
RDI-TV[9]	0.0233	0.0115	0.0031	0.0599	0.0333	0.0274	0.0484	0.0294	0.0267	0.0263	0.0262
Shin[53]	0.0298	0.0136	0.0030	**0.0526**	0.0372	0.0317	0.0465	0.0401	0.0367	0.0269	0.0323
UA[17]	0.0201	0.0142	0.0063	0.0601	0.0520	0.0504	0.0456	0.0310	0.0279	0.0279	0.0273
Proposed method	**0.0189**	**0.0087**	**0.0029**	0.0528	**0.0330**	**0.0237**	**0.0279**	**0.0174**	**0.0158**	**0.0150**	**0.0126**

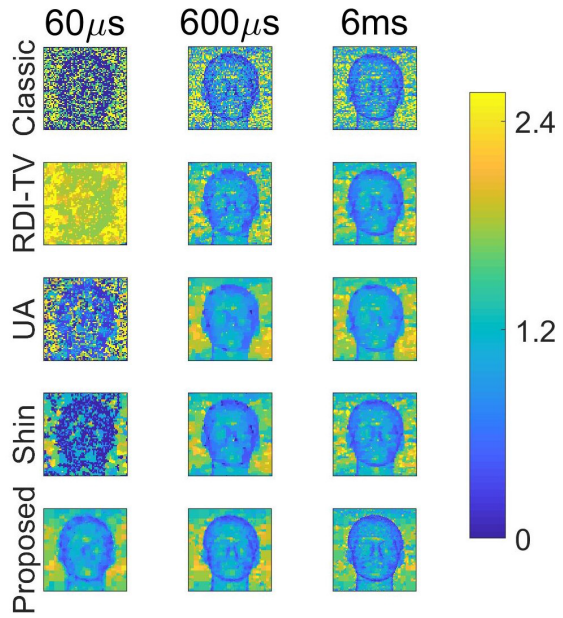

Fig. 16. Reflectivity images obtained by different methods on the mannequin face. Columns from left to right corresponds to data of different acquisition time i.e., 60μs, 600μs and 6ms. Rows from top to bottom corresponds to different methods i.e. cross-correlation, RDI-TV, UA, Shin's algorithm and proposed algorithm.

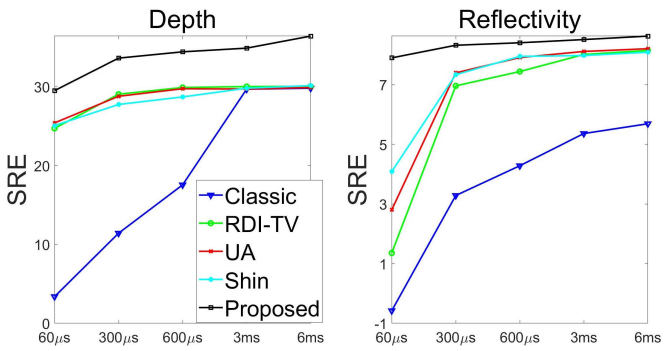

Fig. 17. SRE comparison of depth (left) and reflectivity (right) of different methods, obtained on the mannequin face with 4096 patches/pixels.

and reflectivity images. Fig. 17 shows the SRE results of the compared methods. Thanks to non-uniform sampling, our algorithm delivers better performance than other approaches in both depth and reflectivity. Particularly, the depth images restored by the proposed algorithm in low photon counts reaches similar SRE quality as images of high photon counts restored by other methods (i.e., the SRE level of the proposed algorithm at 60μs is similar to that obtained at 6ms by other methods), which highlights the advantage of the proposed algorithm in presence of few photon and spatial points. RDI-TV and UA show a better level of improvement than Shin's algorithm in depth restoration. The RMSE results are shown in Table III, where the proposed algorithm showed best results in most cases. On the other hand, both Shin and UA algorithms achieve better improvement than RDI-TV in reflectivity restoration. The processing times have also been investigated for all 4 algorithms showing a clear advantage for the proposed algorithm, as highlighted in Table IV.

TABLE IV
PROCESSING TIME (IN SECONDS) OF THE ALGORITHMS COMPARED FOR THE REAL DATA SET FROM THE MANNEQUIN FACE

Acquisition time per pixel (in ms)	0.06	0.3	0.6	3	6
Proposed	**1.67**	**1.27**	**1.28**	**0.96**	**0.68**
RDI-TV[9]	5.60	5.13	3.31	1.85	1.20
Shin[53]	1.82	1.57	1.67	1.32	1.44
UA[17]	2.49	2.05	1.89	3.10	2.92

VII. CONCLUSION

This paper has presented a new formulation for joint restoration of depth and reflectivity images in photon-starved regime. The proposed framework first uses a coarse-to-fine strategy to obtain a good parameter initialization by exploiting information at different resolutions, then learns the spatial correlation by using a non-local graph based approach. A Bayesian approach is used to reconstruct depth and reflectivity profiles. Effort on improving the robustness has been made by combining depth and reflectivity to define the graph nodes. Furthermore, non-uniform sampling was proposed to reduce the computational cost while preserving the useful information. The cost function is efficiently minimized by an alternating direction method of multipliers algorithm. The proposed framework and formulation shows superior performance on synthetic and real data when compared to different methods. Future work will include the study of other regularization terms for depth and reflectivity images, as well as further accelerating the algorithm by adopting parallel computing strategies. The use of adaptive sampling for smart single-photon imaging is interesting and will be the subject of further investigation.

REFERENCES

[1] G. Buller and A. Wallace, "Ranging and three-dimensional imaging using time-correlated single-photon counting and point-by-point acquisition," *IEEE J. Sel. Topics Quantum Electron.*, vol. 13, no. 4, pp. 1006–1015, Jul./Aug. 2007.

[2] A. M. Pawlikowska, A. Halimi, R. A. Lamb, and G. S. Buller, "Single-photon three-dimensional imaging at up to 10 kilometers range," *Opt. Express*, vol. 25, no. 10, pp. 11919–11931, 2017.

[3] R. Tobin *et al.*, "Long-range depth profiling of camouflaged targets using single-photon detection," *Opt. Eng.*, vol. 57, no. 3, 2018, Art. no. 031303.

[4] A. Maccarone *et al.*, "Underwater depth imaging using time-correlated single-photon counting," *Opt. Express*, vol. 23, no. 26, pp. 33911–33926, 2015.

[5] A. Halimi, A. Maccarone, A. McCarthy, S. McLaughlin, and G. S. Buller, "Object depth profile and reflectivity restoration from sparse single-photon data acquired in underwater environments," *IEEE Trans. Comput. Imag.*, vol. 3, no. 3, pp. 472–484, Sep. 2017.

[6] R. Tobin, A. Halimi, A. McCarthy, M. Laurenzis, F. Christnacher, and G. S. Buller, "Three-dimensional single-photon imaging through obscurants," *Opt. Express*, vol. 27, no. 4, pp. 4590–4611, Feb. 2019.

[7] A. Halimi, P. Ciuciu, A. McCarthy, S. McLaughlin, and G. S. Buller, "Fast adaptive scene sampling for single-photon 3D lidar images," in *Proc. IEEE 7th Int. Workshop Comput. Adv. Multi-Sensor Adapt. Process. (CAMSAP)*, Guadeloupe, West Indies, 2019.

[8] P. Caramazza *et al.*, "Neural network identification of people hidden from view with a single-pixel, single-photon detector," *Sci. Rep.*, vol. 8, no. 1, pp. 11945–11950, 2018.

[9] A. Halimi *et al.*, "Restoration of intensity and depth images constructed by using sparse single-photon data," in *Proc. Eur. Signal Process. Conf. (EUSIPCO)*, 2016, pp. 86–90.

[10] A. Halimi, R. Tobin, A. McCarthy, S. McLaughlin, and G. S. Buller, "Restoration of mutltilayered single-photon 3D Lidar images," in *Proc. Eur. Signal Process. Conf. (EUSIPCO)*, 2017, pp. 708–712.

[11] Y. Altmann, A. Maccarone, A. McCarthy, G. Buller, and S. McLaughlin, "Joint spectral clustering and range estimation for 3D scene reconstruction using multispectral Lidar waveforms," in *Proc. Eur. Signal Process. Conf. (EUSIPCO)*, 2016, pp. 513–517.

[12] A. McCarthy *et al.*, "Kilometer-range depth imaging at 1550 nm wavelength using an InGaAs/InP single-photon avalanche diode detector," *Opt. Express*, vol. 21, no. 19, pp. 22098–22113, Sep. 2013.

[13] X. Ren *et al.*, "High-resolution depth profiling using a range-gated CMOS SPAD quanta image sensor," *Opt. Express*, vol. 26, no. 5, pp. 5541–5557, Mar. 2018.

[14] A. M. Wallace, J. Ye, N. Krichel, A. McCarthy, R. Collins, and G. S. Buller, "Full waveform analysis for long-range 3D imaging laser radar," *EURASIP J. Adv. Signal Process.*, vol. 2010, Dec. 2010, Art. no. 896708.

[15] Y. Altmann, X. Ren, A. McCarthy, G. S. Buller, and S. McLaughlin, "Lidar waveform-based analysis of depth images constructed using sparse single-photon data," *IEEE Trans. Image Process.*, vol. 25, no. 5, pp. 1935–1946, May 2016.

[16] D. Shin, A. Kirmani, V. Goyal, and J. Shapiro, "Computational 3D and reflectivity imaging with high photon efficiency," in *Proc. IEEE Int. Conf. Image Process. (ICIP)*, Oct. 2014, pp. 46–50.

[17] J. Rapp and V. K. Goyal, "A few photons among many: Unmixing signal and noise for photon-efficient active imaging," *IEEE Trans. Comput. Imag.*, vol. 3, no. 3, pp. 445–459, Sep. 2017.

[18] A. Kirmani *et al.*, "First-photon imaging," *Science*, vol. 343, no. 6166, pp. 58–61, 2014.

[19] J. Salmon, Z. Harmany, C.-A. Deledalle, and R. Willett, "Poisson noise reduction with non-local PCA," *J. Math. Imag. Vis.*, vol. 48, no. 2, pp. 279–294, 2014.

[20] D. Shin, F. Xu, F. N. Wong, J. H. Shapiro, and V. K. Goyal, "Computational multi-depth single-photon imaging," *Opt. Express*, vol. 24, no. 3, pp. 1873–1888, 2016.

[21] L. Azzari and A. Foi, "Variance stabilization in Poisson image deblurring," in *Proc. IEEE Int. Sym. Biomed. Imag.*, Apr. 2017, pp. 728–731.

[22] M. Makitalo and A. Foi, "Optimal inversion of the generalized Anscombe transformation for Poisson-Gaussian noise," *IEEE Trans. Image Process.*, vol. 22, no. 1, pp. 91–103, Jan. 2013.

[23] Z. T. Harmany, R. F. Marcia, and R. M. Willett, "This is SPIRAL-TAP: Sparse Poisson intensity reconstruction algorithms—Theory and practice," *IEEE Trans. Image Process.*, vol. 21, no. 3, pp. 1084–1096, Mar. 2012.

[24] S. Lefkimmiatis and M. Unser, "Poisson image reconstruction with Hessian Schatten-norm regularization," *IEEE Trans. Image Process.*, vol. 22, no. 11, pp. 4314–4327, Nov. 2013.

[25] J. Li, F. Luisier, and T. Blu, "PURE-LET image deconvolution," *IEEE Trans. Image Process.*, vol. 27, no. 1, pp. 92–105, Jan. 2018.

[26] D. B. Lindell, M. O'Toole, and G. Wetzstein, "Single-photon 3D imaging with deep sensor fusion," *ACM Trans. Graph*, vol. 37, no. 4, Aug. 2018, Art. no. 113.

[27] G. Satat, M. Tancik, O. Gupta, B. Heshmat, and R. Raskar, "Object classification through scattering media with deep learning on time resolved measurement," *Opt. Express*, vol. 25, no. 15, pp. 17466–17479, Jul. 2017.

[28] M. A. T. Figueiredo and R. D. Nowak, "An EM algorithm for wavelet-based image restoration," *IEEE Trans. Image Process.*, vol. 12, no. 8, pp. 906–916, Aug. 2003.

[29] B. Zhang, J. M. Fadili, and J.-L. Starck, "Wavelets, ridgelets, and curvelets for Poisson noise removal," *IEEE Trans. Image Process.*, vol. 17, no. 7, pp. 1093–1108, Jul. 2008.

[30] R. Yan, L. Shao, and Y. Liu, "Nonlocal hierarchical dictionary learning using wavelets for image denoising," *IEEE Trans. Image Process.*, vol. 22, no. 12, pp. 4689–4698, Dec. 2013.

[31] L. I. Rudin, S. Osher, and E. Fatemi, "Nonlinear total variation based noise removal algorithms," *Phys. D, Nonlinear Phenomena*, vol. 60, nos. 1–4, pp. 259–268, 1992.

[32] S. Geman and D. Geman, "Stochastic relaxation, Gibbs distributions, and the Bayesian restoration of images," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. PAMI-6, no. 6, pp. 721–741, Nov. 1984.

[33] K. Dabov, A. Foi, V. Katkovnik, and K. Egiazarian, "Image denoising by sparse 3-D transform-domain collaborative filtering," *IEEE Trans. Image Process.*, vol. 16, no. 8, pp. 2080–2095, Aug. 2007.

[34] A. Buades, B. Coll, and J.-M. Morel, "A review of image denoising algorithms, with a new one," *Multiscale Model. Simul.*, vol. 4, no. 2, pp. 490–530, 2005.

[35] M. Van den Bergh, D. Carton, and L. Van Gool, "Depth SEEDS: Recovering incomplete depth data using superpixels," in *Proc. IEEE Workshop Appl. Comput. Vis. (WACV)*, Jan. 2013, pp. 363–368.

[36] Y. Liu, J. Bioucas-Dias, J. Li, and A. Plaza, "Multi-superpixelization-based convex formulation for joint classification of hyperspectral and lidar data," in *Proc. IEEE Int. Conf. Geosci. Remote Sens. (IGARSS)*, Jul. 2017, pp. 807–810.

[37] B. Mohar and Y. Alavi, "The laplacian spectrum of graphs," *Graph Theory, Combinatorics, Appl.*, vol. 57, no. 3, pp. 871–898, Feb. 1991.

[38] R. Ammanouil, A. Ferrari, and C. Richard, "A graph laplacian regularization for hyperspectral data unmixing," in *Proc. IEEE Conf. Acoust., Speech Signal Process. (ICASSP)*, Apr. 2015, pp. 1637–1641.

[39] S. Boyd, N. Parikh, E. Chu, B. Peleato, and J. Eckstein, "Distributed optimization and statistical learning via the alternating direction method of multipliers," *Found. Trends Mach. Learn.*, vol. 3, no. 1, pp. 1–122, Jan. 2011.

[40] M. A. T. Figueiredo and J. M. Bioucas-Dias, "Restoration of Poissonian images using alternating direction optimization," *IEEE Trans. Image Process.*, vol. 19, no. 12, pp. 3133–3145, Dec. 2010.

[41] M. V. Afonso, J. M. Bioucas-Dias, and M. A. T. Figueiredo, "An augmented Lagrangian approach to the constrained optimization formulation of imaging inverse problems," *IEEE Trans. Image Process.*, vol. 20, no. 3, pp. 681–695, Mar. 2011.

[42] S. Hernandez-Marin, A. M. Wallace, and G. J. Gibson, "Bayesian analysis of lidar signals with multiple returns," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 29, no. 12, pp. 2170–2180, Dec. 2007.

[43] W. Marais and R. Willett, "Proximal-gradient methods for poisson image reconstruction with BM3D-based regularization," in *Proc. IEEE 7th Int. Workshop Comput. Adv. Multi-Sensor Adapt. Process. (CAMSAP)*, Dec. 2017, pp. 183–187.

[44] J. Liu, W. Chi, M. Danilevsky, and J. Han, "Large-scale spectral clustering on graphs," in *Proc. Int. Joint Conf. Artif. Intell.*, 2013, pp. 1486–1492.

[45] P. A. Traganitis and G. B. Giannakis, "Sketched subspace clustering," *IEEE Trans. Signal Process.*, vol. 66, no. 7, pp. 1663–1675, Apr. 2018.

[46] J. Shun, F. Roosta-Khorasani, K. Fountoulakis, and M. W. Mahoney, "Parallel local graph clustering," Apr. 2016, *arXiv:1604.07515*. [Online]. Available: https://arxiv.org/abs/1604.07515

[47] L. Liu, F. Nie, A. Wiliem, Z. Li, T. Zhang, and B. C. Lovell, "Multi-modal joint clustering with application for unsupervised attribute discovery," *IEEE Trans. Image Process.*, vol. 27, no. 9, pp. 4345–4356, Sep. 2018.

[48] Y. W. Y. Andrew Ng and M. I. Jordan, "On spectral clustering: Analysis and an algorithm," in *Proc. Adv. Neural Inf. Process. Syst.*, 2002, pp. 849–856.

[49] A. Halimi *et al.*, "Restoration of depth and intensity images using a graph laplacian regularization," in *Proc. IEEE 7th Int. Workshop Comput. Adv. Multi-Sensor Adapt. Process. (CAMSAP)*, Dec. 2017, pp. 1–5.
[50] A. Halimi, C. Mailhes, J.-Y. Tourneret, and H. Snoussi, "Bayesian estimation of smooth altimetric parameters: Application to conventional and delay/Doppler altimetry," *IEEE Trans. Geosci. Remote Sens.*, vol. 54, no. 4, pp. 2207–2219, Apr. 2016.
[51] A. Halimi, J. M. Bioucas-Dias, N. Dobigeon, G. S. Buller, and S. McLaughlin, "Fast hyperspectral unmixing in presence of nonlinearity or mismodeling effects," *IEEE Trans. Comput. Imag.*, vol. 3, no. 2, pp. 146–159, Jun. 2017.
[52] M.-D. Iordache, J. M. Bioucas-Dias, and A. Plaza, "Collaborative sparse regression for hyperspectral unmixing," *IEEE Trans. Geosci. Remote Sens.*, vol. 52, no. 1, pp. 341–354, Jan. 2014.
[53] D. Shin, A. Kirmani, V. K. Goyal, and J. H. Shapiro, "Photon-efficient computational 3-D and reflectivity imaging with single-photon detectors," *IEEE Trans. Comput. Imag.*, vol. 1, no. 2, pp. 112–125, Jun. 2015.
[54] D. Scharstein *et al.*, "High-resolution stereo datasets with subpixel-accurate ground truth," in *Proc. German Conf. Pattern Recognit. (GCPR)*, 2014, pp. 31–42.

Songmao Chen (S'16) received the B.Eng. degree in artificial intelligence from the Xidian University, China, in 2014. The same year, he began studying at University of Chinese Academy of Science as a postgraduate, majoring in signal and information processing. From 2016, he has been working toward the Ph.D. degree in the same university and conducting research in Xi'an Institute of Optics and Precision Mechanics, Chinese Academy of Science. During 2017–2018, he was sponsored by China Scholarship Council (CSC) as a joint training Ph.D. student in the Single-Photon Group, Heriot-Watt University. He is a student member of the Optical Society of America (OSA) since 2016. His research interest includes single-photon imaging system, image processing and Bayesian inverse problems.

Abderrahim Halimi (S'11–M'14) received the Engineering degree in electronics from the National Polytechnic School of Algiers, Algeria, in 2009, and the M.Sc. and Ph.D. degrees in signal processing from Institut National Polytechnique de Toulouse, Toulouse, France, in 2010 and 2013, respectively. He was a Postdoctoral Research Associate from 2013 to 2018. He is currently an Assistant Professor and a Royal Academy of Engineering (RAEng) Research Fellow with the School of Engineering and Physical Sciences, Heriot-Watt University. His research activities focus on statistical signal and image processing, with a particular interest in Bayesian inverse problems with applications to remote sensing (hyperspectral imaging, satellite altimetry), single-photon depth imaging, and medical imaging.

Ximing Ren received the B.Eng. degree in optical engineering from the Beijing Institute of Technology, China, in 2008, the M.Eng. degree in electrical engineering from Beihang University, China, in 2011, and the Ph.D. degree in physics from Heriot-Watt University, Edinburgh, U.K., in 2015. He carried on three-year Postdoctoral Research on single-photon depth imaging under extreme environments with Single-Photon Group, Heriot-Watt University, under the guidance of Prof. G. S. Buller. He is currently an Assistant Research Professor with the Micron School of Materials Science and Engineering, Boise State University, Boise, ID, USA, and conducts time-resolved single-photon imaging for the characterization of DNA nanostructures. He is a member of The Optical Society (OSA). He provides peer-reviews for the *Transactions on Biomedical Engineering*, *Photonics Journal*, the IEEE PHOTONICS TECHNOLOGY LETTERS, *Optica*, *Optics Express*, *Optics Letters* (OSA), *Optical Engineering* (SPIE), and *Review of Scientific Instruments* (AIP).

Aongus McCarthy (M'03) received the B.Sc. degree from University College Galway, Galway, Ireland, in 1989, the Diploma degree in electronics engineering from the Institute of Technology, Carlow, Ireland, in 1990, the B.Sc. degree in physical optoelectronics from Essex University, Essex, U.K., in 1991, and the Ph.D. degree in physics from Heriot-Watt University, Edinburgh, U.K., in 2002. He worked in industry as a Design Team Leader on the development of a thermal transfer printing system from 1992 to 1997. He is currently a Research Fellow with the School of Engineering and Physical Sciences, Heriot-Watt University. His research interests include optical and optomechanical system design, time-of-flight depth imaging, single-photon counting technologies, and microscope systems. He is a member of the Optical Society of America and the IEEE Photonics Society.

Xiuqin Su received the B.Eng. degree in radio techniques from Xian Jiaotong University and the M.Eng. degree in optical instrument from the Xi'an Institute of Optics and Precision Mechanics (XIOPM), CAS, in 1986 and 1989, respectively. From 1998 to 1999, she was a Visiting Scholar with the National Research Council of Canada, Information and Technology Group. She is currently a Professor with XIOPM, CAS, and the University of Chinese Academy of Science (UCAS). Her research interest includes single-photon imaging, non-line-of-sight imaging, and lensless imaging.

Stephen McLaughlin (F'11) was born in Clydebank, Scotland, in 1960. He received the B.Sc. degree in electronics and electrical engineering from the University of Glasgow in 1981 and the Ph.D. degree from the University of Edinburgh in 1990. From 1981 to 1984, he was a Development Engineer in industry. In 1986, he joined the University of Edinburgh, where he studied the performance of linear adaptive algorithms in high noise and nonstationary environments. In 1988, he joined the Academic Staff at Edinburgh, and from 1991 to 2001, he held a Royal Society University Research Fellowship to study nonlinear signal processing techniques. In 2002, he was awarded the personal Chair in electronic communication systems at the University of Edinburgh. In October 2011, he joined Heriot-Watt University as a Professor of signal processing and the Head of the School of Engineering and Physical Sciences. His research interests lie in the fields of adaptive signal processing and nonlinear dynamical systems theory and their applications to biomedical, energy and communication systems. He is a Fellow of the Royal Academy of Engineering, the Royal Society of Edinburgh, and the Institute of Engineering and Technology.

Gerald S. Buller received the B.Sc. degree (Hons.) in natural philosophy from the University of Glasgow, Glasgow, U.K., in 1986, and the Ph.D. degree in physics from Heriot-Watt University, Edinburgh, U.K., in 1989. In 2002, he co-founded Helia Photonics Ltd., Livingston, U.K. He is currently a Professor of physics with Heriot-Watt University. His current research interests include single-photon detection methods, particularly at short-wave infrared wavelengths, and applications involving the use of single photons. These applications include single-photon depth imaging and novel quantum communications protocols. He is a Fellow of the Royal Society of Edinburgh, the U.K. Institute of Physics, and the Optical Society of America. In 2015, he received the EPSRC Established Career Fellowship in Quantum Technology to research sparse photon imaging.